OIPE

RAW SEQUENCE LISTING   DATE: 11/06/2001
PATENT APPLICATION: US/09/884,889   TIME: 15:16:37

Input Set : A:\DIVER1100-4.ST25.txt
Output Set: N:\CRF3\11062001\I884889.raw

```
 2 <110> APPLICANT: DIVERSA CORPORATION
 3       ROBERTSON, Dan
 4       SANYAL, Indrajit
 5       ADHIKARI, Robert
 7 <120> TITLE OF INVENTION: CATALASES
 9 <130> FILE REFERENCE: DIVER1100-4
11 <140> CURRENT APPLICATION NUMBER: US 09/884,889
12 <141> CURRENT FILING DATE: 2001-06-19
14 <150> PRIOR APPLICATION NUMBER: US 09/412,347
15 <151> PRIOR FILING DATE: 1999-10-05
17 <150> PRIOR APPLICATION NUMBER: US 08/951,844
18 <151> PRIOR FILING DATE: 1997-10-16
20 <150> PRIOR APPLICATION NUMBER: US 08/674,887
21 <151> PRIOR FILING DATE: 1996-07-03
23 <160> NUMBER OF SEQ ID NOS: 8
25 <170> SOFTWARE: PatentIn version 3.0
27 <210> SEQ ID NO: 1
28 <211> LENGTH: 52
29 <212> TYPE: DNA
30 <213> ORGANISM: Artificial sequence
32 <220> FEATURE:
33 <223> OTHER INFORMATION: Primer for PCR
35 <400> SEQUENCE: 1
36 ccgagaattc attaaagagg agaaattaac tatgaataac gcatccgctg ac          52
39 <210> SEQ ID NO: 2
40 <211> LENGTH: 31
41 <212> TYPE: DNA
42 <213> ORGANISM: Artificial sequence
44 <220> FEATURE:
45 <223> OTHER INFORMATION: Primer for PCR
47 <400> SEQUENCE: 2
48 gcaaagctgc agcgcagcat tttcgaaagg c                                 31
51 <210> SEQ ID NO: 3
52 <211> LENGTH: 52
53 <212> TYPE: DNA
54 <213> ORGANISM: Artificial sequence
56 <220> FEATURE:
57 <223> OTHER INFORMATION: Primer for PCR
59 <400> SEQUENCE: 3
60 ccgagaattc attaaagagg agaaattaac tatggaaaat cacaaacact ca          52
63 <210> SEQ ID NO: 4
64 <211> LENGTH: 31
65 <212> TYPE: DNA
66 <213> ORGANISM: Artificial sequence
68 <220> FEATURE:
69 <223> OTHER INFORMATION: Primer for PCR
71 <400> SEQUENCE: 4
```

ENTERED

```
 72 ctggccaaac tagactttat tccatggaag c                               31
 75 <210> SEQ ID NO: 5
 76 <211> LENGTH: 2262
 77 <212> TYPE: DNA
 78 <213> ORGANISM: Alcaligenes (Deleya) aquamarinus
 80 <400> SEQUENCE: 5
 81 atgaataacg catccgctga cgatctacac agtagcttgc agcaaagatg cagagcattt    60
 83 gttcccttgg tatcgccaag gcatagagca ataagggaga gagctatgag cggtaaatgt   120
 85 cctgtcatgc acggtggtaa cacctcgacc ggtacttcca acaaagattg gtggccggaa   180
 87 gggttgaacc tggatatttt gcatcagcaa gatcgcaaat cagacccgat ggatccggat   240
 89 ttcaactacc gtgaagaagt acgcaagctc gatttcgacg cgctgaagaa agatgtccac   300
 91 gcgttgatga ccgatagcca agagtggtgg cccgctgact ggggggcacta cggcggtttg   360
 93 atgatccgta tggcttggca ctccgctggc acctaccgta ttgctgatgg ccgtgggggc   420
 95 ggtggtaccg gaagccagcg ctttgcaccg ctcaactcct ggccggacaa cgtcagcctg   480
 97 gataaagcgc gccgtctgct gtggccgatc aagaagaagt acggcaacaa aatcagctgg   540
 99 gcagacctga tgattctggc tggcaccgtg gcttatgagt ccatgggctt acctgcttac   600
101 ggcttctctt tcggccgcgt cgatatttgg gaacccgaaa aagatatcta ctggggtgac   660
103 gaaaaagagt ggctggcacc ttctgacgaa cgctacggcg acgtgaacaa gccagagacc   720
105 atggaaaacc cgctggcggc tgtccaaatg ggtctgatct atgtgaaccc ggaaggtgtt   780
107 aacggccacc ctgatccgct gagaaccgca cagcaggtac ttgaaacctt cgcccgtatg   840
109 gcgatgaacg acgaaaaaac cgcagccctc acagctggcg gccacaccgt cggtaattgt   900
111 cacggtaatg gcaatgcctc tgcgttagcc cctgacccaa aagcctctga cgttgaaaac   960
113 cagggcttag gttggggcaa ccccaacatg cagggcaagg caagcaacgc cgtgacctcg  1020
115 ggtatcgaag gtgcttggac caccaacccc acgaaattcg atatgggcta tttcgacctg  1080
117 ctgttcggct acaattggga actgaaaaag agtcctgccg gtgcccacca ttgggaaccg  1140
119 attgacatca aaaaggaaaa caagccggtt gacgccagcg acccctctat tcgccacaac  1200
121 ccgatcatga ccgatgcgga tatggcgata aaggtaaatc cgacctatcg cgctatctgc  1260
123 gaaaaattca tggccgatcc tgagtacttc aagaaaactt tcgcgaaggc gtggttcaag  1320
125 ctgacgcacc gtgacctggg cccgaaatca cgttacatcg gcccggaagt gccggcagaa  1380
127 gacctgattt ggcaagaccc gattccggca ggtaacaccg actactgcga agaagtggtc  1440
129 aagcagaaaa ttgcacaaag tggcctgagc attagtgaga tggtctccac cgcttgggac  1500
131 agtgcccgta cttatcgcgg ttccgatatg cgcggcggtg ctaacggtgc ccgcattcgc  1560
133 ttggccccac agaacgagtg gcaggcaac gagccggagc gcctggcgaa agtgctgagc   1620
135 gtctacgagc agatctctgc cgacaccggc gctagcatcg cggacgtgat cgttctggcc  1680
137 ggtagcgtag gcatcgagaa agccgcgaaa gcagcaggtt acgatgtgcg cgttcccttc  1740
139 ctgaaaggcc gtggcgatgc gaccgccgag atgaccgacg cagactcctt cgcaccgctg  1800
141 gagccgctgg ccgatggctt ccgcaactgg cagaagaaag agtatgtggt gaagccggaa  1860
143 gagatgctgc tggatcgtgc gcagctgatg ggcttaaccg gcccggaaat gaccgtgctg  1920
145 ctgggcggta tgcgcgtact gggcaccaac tatggtggca ccaaacacgg cgtattcacc  1980
147 gattgtgaag gccagttgac caacgacttt tttgtgaacc tgaccgatat ggggaacagc  2040
149 tggaagccgg taggtagcaa cgcctacgaa atccgcgacc gcaagaccgg tgccgtgaag  2100
151 tggaccgcct cgcgggtgga tctggtattt ggttccaact cgctactgcg ctcttacgca  2160
153 gaagtgtacg cccaggacga taacggcgag aagttcgtca gagacttcgt cgccgcctgg  2220
155 accaaagtga tgaacgccga ccgtttcgac gtcgcgtcgt aa                     2262
158 <210> SEQ ID NO: 6
159 <211> LENGTH: 753
160 <212> TYPE: PRT
161 <213> ORGANISM: Alcaligenes (Deleya) aquamarinus
163 <400> SEQUENCE: 6
```

RAW SEQUENCE LISTING   DATE: 11/06/2001
PATENT APPLICATION: US/09/884,889   TIME: 15:16:37

Input Set : A:\DIVER1100-4.ST25.txt
Output Set: N:\CRF3\11062001\I884889.raw

165 Met Asn Asn Ala Ser Ala Asp Asp Leu His Ser Ser Leu Gln Gln Arg
166 1   5   10   15
168 Cys Arg Ala Phe Val Pro Leu Val Ser Pro Arg His Arg Ala Ile Arg
169   20   25   30
171 Glu Arg Ala Met Ser Gly Lys Cys Pro Val Met His Gly Gly Asn Thr
172   35   40   45
174 Ser Thr Gly Thr Ser Asn Lys Asp Trp Trp Pro Glu Gly Leu Asn Leu
175   50   55   60
177 Asp Ile Leu His Gln Gln Asp Arg Lys Ser Asp Pro Met Asp Pro Asp
178 65   70   75   80
180 Phe Asn Tyr Arg Glu Glu Val Arg Lys Leu Asp Phe Asp Ala Leu Lys
181   85   90   95
183 Lys Asp Val His Ala Leu Met Thr Asp Ser Gln Glu Trp Trp Pro Ala
184   100   105   110
186 Asp Trp Gly His Tyr Gly Gly Leu Met Ile Arg Met Ala Trp His Ser
187   115   120   125
189 Ala Gly Thr Tyr Arg Ile Ala Asp Gly Arg Gly Gly Gly Gly Thr Gly
190   130   135   140
192 Ser Gln Arg Phe Ala Pro Leu Asn Ser Trp Pro Asp Asn Val Ser Leu
193 145   150   155   160
195 Asp Lys Ala Arg Arg Leu Leu Trp Pro Ile Lys Lys Lys Tyr Gly Asn
196   165   170   175
198 Lys Ile Ser Trp Ala Asp Leu Met Ile Leu Ala Gly Thr Val Ala Tyr
199   180   185   190
201 Glu Ser Met Gly Leu Pro Ala Tyr Gly Phe Ser Phe Gly Arg Val Asp
202   195   200   205
204 Ile Trp Glu Pro Glu Lys Asp Ile Tyr Trp Gly Asp Glu Lys Glu Trp
205   210   215   220
207 Leu Ala Pro Ser Asp Glu Arg Tyr Gly Asp Val Asn Lys Pro Glu Thr
208 225   230   235   240
210 Met Glu Asn Pro Leu Ala Ala Val Gln Met Gly Leu Ile Tyr Val Asn
211   245   250   255
213 Pro Glu Gly Val Asn Gly His Pro Asp Pro Leu Arg Thr Ala Gln Gln
214   260   265   270
216 Val Leu Glu Thr Phe Ala Arg Met Ala Met Asn Asp Glu Lys Thr Ala
217   275   280   285
219 Ala Leu Thr Ala Gly Gly His Thr Val Gly Asn Cys His Gly Asn Gly
220   290   295   300
222 Asn Ala Ser Ala Leu Ala Pro Asp Pro Lys Ala Ser Asp Val Glu Asn
223 305   310   315   320
225 Gln Gly Leu Gly Trp Gly Asn Pro Asn Met Gln Gly Lys Ala Ser Asn
226   325   330   335
228 Ala Val Thr Ser Gly Ile Glu Gly Ala Trp Thr Thr Asn Pro Thr Lys
229   340   345   350
231 Phe Asp Met Gly Tyr Phe Asp Leu Leu Phe Gly Tyr Asn Trp Glu Leu
232   355   360   365
234 Lys Lys Ser Pro Ala Gly Ala His His Trp Glu Pro Ile Asp Ile Lys
235   370   375   380
237 Lys Glu Asn Lys Pro Val Asp Ala Ser Asp Pro Ser Ile Arg His Asn

```
238 385                     390                     395                     400
240 Pro Ile Met Thr Asp Ala Asp Met Ala Ile Lys Val Asn Pro Thr Tyr
241              405                     410                     415
243 Arg Ala Ile Cys Glu Lys Phe Met Ala Asp Pro Glu Tyr Phe Lys Lys
244              420                     425                     430
246 Thr Phe Ala Lys Ala Trp Phe Lys Leu Thr His Arg Asp Leu Gly Pro
247          435                     440                     445
249 Lys Ser Arg Tyr Ile Gly Pro Glu Val Pro Ala Glu Asp Leu Ile Trp
250      450                     455                     460
252 Gln Asp Pro Ile Pro Ala Gly Asn Thr Asp Tyr Cys Glu Glu Val Val
253 465                     470                     475                     480
255 Lys Gln Lys Ile Ala Gln Ser Gly Leu Ser Ile Ser Glu Met Val Ser
256              485                     490                     495
258 Thr Ala Trp Asp Ser Ala Arg Thr Tyr Arg Gly Ser Asp Met Arg Gly
259              500                     505                     510
261 Gly Ala Asn Gly Ala Arg Ile Arg Leu Ala Pro Gln Asn Glu Trp Gln
262          515                     520                     525
264 Gly Asn Glu Pro Glu Arg Leu Ala Lys Val Leu Ser Val Tyr Glu Gln
265      530                     535                     540
267 Ile Ser Ala Asp Thr Gly Ala Ser Ile Ala Asp Val Ile Val Leu Ala
268 545                     550                     555                     560
270 Gly Ser Val Gly Ile Glu Lys Ala Ala Lys Ala Ala Gly Tyr Asp Val
271              565                     570                     575
273 Arg Val Pro Phe Leu Lys Gly Arg Gly Asp Ala Thr Ala Glu Met Thr
274          580                     585                     590
276 Asp Ala Asp Ser Phe Ala Pro Leu Glu Pro Leu Ala Asp Gly Phe Arg
277          595                     600                     605
279 Asn Trp Gln Lys Lys Glu Tyr Val Val Lys Pro Glu Glu Met Leu Leu
280      610                     615                     620
282 Asp Arg Ala Gln Leu Met Gly Leu Thr Gly Pro Glu Met Thr Val Leu
283 625                     630                     635                     640
285 Leu Gly Gly Met Arg Val Leu Gly Thr Asn Tyr Gly Gly Thr Lys His
286          645                     650                     655
288 Gly Val Phe Thr Asp Cys Glu Gly Gln Leu Thr Asn Asp Phe Phe Val
289          660                     665                     670
291 Asn Leu Thr Asp Met Gly Asn Ser Trp Lys Pro Val Gly Ser Asn Ala
292      675                     680                     685
294 Tyr Glu Ile Arg Asp Arg Lys Thr Gly Ala Val Lys Trp Thr Ala Ser
295      690                     695                     700
297 Arg Val Asp Leu Val Phe Gly Ser Asn Ser Leu Leu Arg Ser Tyr Ala
298 705                     710                     715                     720
300 Glu Val Tyr Ala Gln Asp Asp Asn Gly Glu Lys Phe Val Arg Asp Phe
301              725                     730                     735
303 Val Ala Ala Trp Thr Lys Val Met Asn Ala Asp Arg Phe Asp Val Ala
304              740                     745                     750
306 Ser
309 <210> SEQ ID NO: 7
310 <211> LENGTH: 2238
311 <212> TYPE: DNA
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/884,889

DATE: 11/06/2001
TIME: 15:16:37

Input Set : A:\DIVER1100-4.ST25.txt
Output Set: N:\CRF3\11062001\I884889..raw

```
312 <213> ORGANISM: Microscilla furvescens
314 <400> SEQUENCE: 7
315 atggaaaatc acaaacactc aggatcttct acgtataaca caaacactgg cggaaaatgc    60
317 ccttttaccg gaggttcgct taagcaaagt gcaggtggcg gcaccaaaaa cagggattgg   120
319 tggcccaaca tgctcaacct cggcatctta cgccaacatt catcgctatc ggacccaaac   180
321 gacccggatt ttgactatgc cgaagagttt aagaagctag atctggcagc ggttaaaaag   240
323 gacctggcag cgctaatgac agattcacag gactggtggc cagcagatta cggtcattat   300
325 ggccccttct ttatacgcat ggcgtggcac agcgccggca cctaccgtat cggtgatggc   360
327 cgtggtggcg gtggctccgg ctcacagcgc ttcgcgcctc tcaatagctg gccagacaat   420
329 gccaatctgg ataaagcacg cttgcttctt tggcccatca aacaaaaata cggtcgaaaa   480
331 atctcctggg cggatctaat gatactcaca ggaaacgtag ctctggaaac tatgggcttt   540
333 aaaacttttg gttttgcagg tggcagagca gatgtatggg agcctgaaga agatgtatac   600
335 tggggagcag aaaccgaatg gctgggagac aagcgctatg aaggtgaccg agagctcgaa   660
337 aatcccctgg gagccgtaca aatgggactc atctatgtaa accccgaagg acccaacggc   720
339 aagccagacc ctatcgctgc tgcgcgtgat attcgtgaga cttttggccg aatggcaatg   780
341 aatgacgaag aaaccgtggc tctcatagcg ggtggacaca ccttcggaaa aacccatggt   840
343 gctgccgatg cggagaaata tgtgggccga gagcctgccg ccgcaggtat tgaagaaatg   900
345 agcctggggt ggaaaaacac ctacggcacc ggacacggtg cggataccat caccagtgga   960
347 ctagaaggcg cctggaccaa gacccctact caatggagca ataacttttt tgaaaacctc  1020
349 tttggttacg agtgggagct taccaaaagt ccagctggag cttatcagtg gaaaccaaaa  1080
351 gacggtgccg gggctggcac cataccggat gcacatgatc ccagcaagtc gcacgctcca  1140
353 tttatgctca ctacggacct ggcgctgcgc atggaccctg attacgaaaa aatttctcga  1200
355 cggtactatg aaaaccctga tgagtttgca gatgctttcg cgaaagcatg gtacaaactg  1260
357 acacacagag atatgggacc aaaggtgcgc tacctgggac cagaagtgcc tcaggaagac  1320
359 ctcatctggc aagaccctat accagatgta agccatcctc ttgtagacga aaacgatatt  1380
361 gaaggcctaa aagccaaaat cctggaatcg ggactgacgg taagcgagct ggtaagcacg  1440
363 gcatgggctt ctgcatctac tttttagaaac tctgacaagc gcggcggtgc caacggtgca  1500
365 cgtatacgac tggcccccaca aaaagactgg gaagtaaaca accctcagca acttgccagg  1560
367 gtactcaaaa cactagaagg tatccaggag gactttaacc aggcgcaatc agataacaaa  1620
369 gcagtatcgt tggccgacct gattgtgctg ccggctgtg cgggtgtaga aaaagctgca  1680
371 aaagatgctg gccatgaggt gcaggtgcct ttcaacccgg gacgagcgga tgccaccgct  1740
373 gagcaaaccg atgtggaagc tttcgaagca ctagagccag cggctgacgg ctttagaaac  1800
375 tacattaaac cggagcataa agtatccgct gaggaaatgc tcgtagaccg ggcgcagctt  1860
377 ctgtcgcttt cggcaccaga aatgactgct ttggtaggcg gtatgcgtgt actgggcacc  1920
379 aactacgacg gttcgcagca tggagtgttt acaaataagc cgggtcagct atccaatgac  1980
381 ttctttgtaa acctgctaga cctcaacact aaatggcgag ccagcgatga atcagacaaa  2040
383 gtttttgaag cagagactt caaaactggc gaagtaaagt ggagtggcac ccgggtagac  2100
385 ctgatcttcg gatccaattc cgagctaaga gccctcgcag aagtgtacgg ctgtgcagat  2160
387 tctgaagaaa agtttgttaa agattttgtg aaggcctggg ccaaagtaat ggacctggac  2220
389 cggtttgatc tgaaataa                                                 2238
392 <210> SEQ ID NO: 8
393 <211> LENGTH: 745
394 <212> TYPE: PRT
395 <213> ORGANISM: Microscilla furvescens
397 <400> SEQUENCE: 8
399 Met Glu Asn His Lys His Ser Gly Ser Ser Thr Tyr Asn Thr Asn Thr
400 1               5                   10                  15
402 Gly Gly Lys Cys Pro Phe Thr Gly Gly Ser Leu Lys Gln Ser Ala Gly
403             20                  25                  30
```

VERIFICATION SUMMARY
PATENT APPLICATION: **US/09/884,889**

DATE: 11/06/2001
TIME: 15:16:38

Input Set : **A:\DIVER1100-4.ST25.txt**
Output Set: **N:\CRF3\11062001\I884889.raw**